

SYSTEM AND METHOD FOR ACQUIRING AND
PROCESSING COMPLEX IMAGES

CROSS REFERENCE TO RELATED APPLICATIONS

This application claims priority from U.S.
Provisional Patent Application Serial No. 60/410,240,
filed September 12, 2002, and entitled "System and Method
5 of Image Matching and Registration with an Automated
Multi-View Target Searching Mechanism," U.S. Provisional
Patent Application Serial No. 60/410,157, filed September
12, 2002, and entitled "System and Method for Comparing
Holographic Images," Application Serial No. 60/410,153,
10 filed September 12, 2002, and entitled "System and Method
of Aligning Difference Images" and Application Serial No.
60/410,152, filed September 12, 2002, and entitled
"System and Method of Generating a Difference Between
Complex Images."

15

TECHNICAL FIELD OF THE INVENTION

The present invention relates in general to the
field of data processing and more specifically to a
20 system and method for acquiring and processing complex
images.

BACKGROUND OF THE INVENTION

Holograms captured with a digital acquisition system contain information about the material characteristics and topology of the object being viewed. By capturing sequential holograms of different instances of the same object, changes between objects can be measured in several dimensions. Digital processing of the holograms allows for a direct comparison of the actual image waves of the object. These image waves contain significantly more information on small details than conventional non-holographic images, because the image phase information is retained in the holograms, but lost in conventional images. The end goal of a system that compares holographic images is to quantify the differences between objects and determine if a significant difference exists.

The process of comparing holograms is a difficult task because of the variables involved in the hologram generation process and object handling. In particular, in order to effectively compare corresponding holographic images, two or more holographic images must be acquired and registered or "matched" such that the images closely correspond. Additionally, after the holographic images are acquired and registered, the images are compared to determine differences between the images. Existing techniques for registering and comparing corresponding images often requires significant processing and time. Such time and processing requirements limit the throughput and overall efficiency of digital holographic imaging systems.

BRIEF DESCRIPTION OF THE DRAWINGS

A more complete understanding of the present
embodiments and advantages thereof may be acquired by
referring to the following description taken in
5 conjunction with the accompanying drawings, in which like
reference numbers indicate like features, and wherein:

FIGURE 1 is a flow diagram showing an intensity
based registration method;

FIGURE 2 is a flow diagram showing a magnitude based
10 registration method;

FIGURE 3 is a flow diagram showing a registration
method for holographic phase images;

FIGURE 4 is a flow diagram showing a registration
method for holographic complex images;

15 FIGURE 5 is a flow diagram of a simplified
registration system that eliminates the confidence value
computation;

FIGURE 6 is a flow diagram showing a simplified
registration system for holographic complex images;

20 FIGURE 7 is demonstrative diagram of a wafer for
determining positional refinement;

FIGURE 8 is a diagram of a digital holographic
imaging system;

25 FIGURE 9 is an image of a hologram acquired from a
CCD camera;

FIGURE 10 is an enlarged portion of FIGURE 10
showing fringe detail;

FIGURE 11 is a holographic image transformed using a
Fast Fourier Transform (FFT) operation;

30 FIGURE 12 is a holographic image showing a sideband;

FIGURE 13 is a quadrant of a hologram FFT centered at the carrier frequency;

FIGURE 14 shows the sideband of FIGURE 14 after application of a Butterworth lowpass filter;

5 FIGURE 15 shows a magnitude image;

FIGURE 16 shows a phase image;

FIGURE 17 shows a difference image;

FIGURE 18 shows a second difference image;

FIGURE 19 shows a thresholded difference image;

10 FIGURE 20 shows a second thresholded difference image;

FIGURE 21 shows an image of two thresholded difference images following a logical AND operation;

FIGURE 22 shows a magnitude image with defects; and

15 FIGURE 23 shows a phase image with defects.

DETAILED DESCRIPTION

Preferred embodiments and their advantages are best understood by reference to FIGURES 1 through 23, wherein like numbers are used to indicate like and corresponding parts.

The following invention relates to digital holographic imaging systems and applications as described, for instance, in U.S. Patent No. 6,078,392 entitled *Direct-to-Digital Holography and Holography*, U.S. Patent No. 6,525,821 entitled, *Improvements to Acquisition and Replay Systems for Direct to Digital Holography and Holography*, U.S. Patent Application Serial no. 09/949,266 entitled *System and Method for Correlated Noise Removal in Complex Imaging Systems* now issued as U.S. Patent no. _____ and U.S. Patent Application Serial No. 09/949,423 entitled, *System and Method for Registering Complex Images* now issued as U.S. Patent no. _____, all of which are incorporated herein by reference.

The present invention encompasses the automated image registration and processing techniques that have been developed to meet the special needs of Direct-to-Digital Holography (DDH) defect inspection systems as described herein. In DDH systems, streamed holograms may be compared on a pixel-by-pixel basis for defect detection after hologram generation.

One embodiment of the present invention includes systems and methods for automated image matching and registration with a feedback confidence measure are described below. The registration system provides a

techniques and algorithms for multiple image matching tasks in DDH systems, such as runtime wafer inspection, scene matching refinement, and rotational wafer alignment. In some embodiments, a system for
5 implementing this registration system may include several major aspects including: a search strategy, multiple data input capability, normalized correlation implemented in the Fourier domain, noise filtering, correlation peak pattern search, confidence definition and computation,
10 sub-pixel accuracy modeling, and automated target search mechanism.

Image Registration

The Fourier transform of a signal is a unique
15 representation of the signal, i.e. the information contents are uniquely determined by each other in two different domains. Therefore, given two images with some degree of congruence, $f_1(x,y)$ and $f_2(x,y)$, with Fourier transforms, $F1(w_x,w_y)$ and $F2(w_x,w_y)$, their spatial
20 relationship can also be uniquely represented by the relationship between their Fourier transforms. For example, an Affine transformation between two signals in the spatial domain can be represented uniquely by their Fourier transforms based on the shifting theorem, scaling
25 theorem, and rotational theorem of the Fourier transform. If there is an affine transformation between $f_1(x,y)$ and $f_2(x,y)$, their spatial relationship can be expressed as:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$$

where

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

represents rotational, scaling, and skew

5 differences; and $\begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$ represents translations. If it is a noise-free environment, the two images are related to each other by:

$$f_1(x, y) = f_2(ax + by + x_0, cx + dy + y_0) ;$$

10

and their Fourier transforms are related as follows:

$$F_2(w_x, w_y) = |A| F_1 \left(A^T \begin{bmatrix} w_x \\ w_y \end{bmatrix} \right) \cdot e^{-j(w_x x_0 + w_y y_0)} ,$$

15 where A^T denotes the transpose of A and $|A|$ is its determinant. The importance of this derivation is that this equation separates the affine parameters into two groups in the Fourier space: translations and linear transformation, which tells us that the translations are
20 determined by Fourier phase difference while magnitude is shift-invariant and related to each other by the linear component $|A|$.

In the simplest case: translation model, i.e. one image is simply a shifted version of another image, as
25 in:

$$f_1(x, y) = f_2(x + x_0, y + y_0) .$$

Their Fourier transforms have the following relationship:

$$F_2(w_x, w_y) = F_1(w_x, w_y) \cdot e^{j(w_x x_0 + w_y y_0)},$$

- 5 based on the Fourier shift theorem, which is equivalent to:

$$\frac{F_2(w_x, w_y) F_1^*(w_x, w_y)}{|F_2(w_x, w_y) F_1^*(w_x, w_y)|} = e^{j(w_x x_0 + w_y y_0)}.$$

- 10 The left-hand side of the equation above is the cross power spectrum normalized by the maximum power possible of two signals. It is also called coherence function. Two signals have the same magnitude spectra but a linear phase difference corresponding to spatial translations.
- 15 The coherence function of two images, $\Gamma_{12}(w_x, w_y)$, is also related to their cross correlation defined by power spectral densities (PSD) and cross power spectral density (XPSD) by

20
$$\Gamma_{12}(w_x, w_y) = \frac{xpsd}{\sqrt{psd_1 \cdot psd_2}},$$

- where $xpsd$ is the cross power spectral density of the two images, and psd_1 and psd_2 are the power spectral densities of f_1 and f_2 respectively. Assuming it is a stationary
- 25 stochastic process, its true PSD is the Fourier transform of the true autocorrelation function. The Fourier transform of the autocorrelation function of an image provides a sample estimate of the PSD. By the same token,

cross power density $xpsd$ can be estimated by the 2-D Fourier transform of f_2 multiplied by the complex conjugate of the 2-D Fourier transform of f_1 . Therefore, the coherence function of two images may be estimated by

5

$$\Gamma_{12}(w_x, w_y) \cong \frac{F_2(w_x, w_y) F_1^*(w_x, w_y)}{|F_2(w_x, w_y) F_1^*(w_x, w_y)|}$$

The coherence function above is a function of spatial frequency with its magnitude indicating the amplitude of power present in the cross-correlation function. It is also a frequency representation of cross correlation (CC), i.e. the Fourier transform of cross correlation, as indicated by the correlation theorem of the Fourier transform:

15

$$f_1(x, y) \otimes f_2(x, y) \Leftrightarrow F_2(w_x, w_y) F_1(-w_x, -w_y),$$

where \otimes denotes spatial correlation. For real signal, the Fourier transform is conjugate symmetric, i.e.

20

$$F_1(-w_x, -w_y) = F_1^*(w_x, w_y)$$

The maximum correlated power possible is an estimate of $\sqrt{psd1 \cdot psd2}$. The magnitude-squared coherence, $|\Gamma_{12}(w_x, w_y)|^2$, is a real function between 0 and 1 which gives a measure of correlation between the two images at each frequency. At a given frequency, when the correlated power is the same as the maximum correlated power possible, the two images observe the same patterns and the powers only varies by a scale factor. In this case, $CC = 1$. When the two images have different patterns, the powers will be out of phase in the two power spectrum densities and the

30

cross-power spectral density will have lower power than the maximum possible case. For these reasons, the coherence function can be used in image matching and the coherence value is a measure of correlation between the
5 two images.

Based on the theory described above, the matching position of the two images, i.e. point of registration, can be derived by locating where the maximum CC is in the spatial domain. The inverse Fourier transform of CC (i.e.
10 an estimate of the coherent function) is

$$F^{-1}\left(\frac{F_2(w_x, w_y) F_1^*(w_x, w_y)}{|F_2(w_x, w_y) F_1^*(w_x, w_y)|}\right) = F^{-1}(e^{j(w_x x_0 + w_y y_0)}) = \delta(x_0, y_0),$$

which is a Dirac delta function. This is the
15 representation of CC in the spatial domain and the position of the delta function is exactly where the registration is located.

For real signal and system with limited bandwidth (finite size of discrete Fourier transform) and
20 assumption of periodic extension of spatial signal, the delta function becomes a unit pulse. Given two signals with some degree of congruence, signal power in their cross power spectrum is mostly concentrated in a coherent peak in the spatial domain, located at the point of
25 registration. Noise power is distributed randomly in some coherent peaks. The amplitude of the coherent peak is a direct measure of the congruence between the two images. More precisely, the power in the coherent peak corresponds to the percentage of overlapping areas, while

the power in incoherent peaks correspond to the percentage of non-overlapping areas.

Effect of Noise, Feature Space Selection, and Filters

5 Ideally, the coherence of the features of interest should be 1 at all frequencies in the frequency domain and a delta pulse at the point of registration in the spatial domain. However, noise will typically distort the correlation surface. These noises include time-varying
10 noise (A/C noise) such as back-reflection noise, carrier drifting, and variation caused by process change, fixed-pattern noise (D/C noise) such as illumination non-uniformity, bad pixels, camera scratch, dusts on optical path, and focus difference, and stage tilting; and (3)
15 random noise.

 If these noises are present, one may think of one image as a superposition of three images in both additive and multiplicative ways:

20
$$f_n(x,y) = N_m(x,y) * f(x,y) + N_a(x,y),$$

 where $N_m(x,y)$ is multiplicative noise source; $N_a(x,y)$ is an additive noise source; and $f_n(x,y)$ is the signal distorted by noise.

25

$$F_n(x,y) = F_m(x,y) \otimes F(x,y) + F_a(x,y)$$

 where $F_m(x,y)$ is the Fourier transform of the multiplicative noise source; $F_a(x,y)$ is the Fourier

transform of the additive noise source; and $F_n(x,y)$ is the Fourier transform the signal distorted by noise.

The observed signal is $f_n(x,y)$ with its Fourier transform $F_n(x,y)$. The objective of noise processing is to
5 make the coherent peak converge on the signal only. There are primarily two ways to achieve this goal: (1) to reconstruct its original signal $f(x,y)$ or its original Fourier transform $F(x,y)$ from the observed signal; (2) to reduce the noise as much as possible to increase the
10 probability of convergence on the signal even the signal is partially removed or attenuated.

The first method of noise removal requires noise modeling with each noise source typically requiring a different model. The second method focuses on noise
15 removal by any means even it also removes or attenuates the signal, which gives us much more room to operate. Therefore, we mainly use the second technique for the task of image matching. Furthermore, it is beneficial to think of the issue in both spatial domain and frequency
20 domain. The observations below have been considered in the design of noise resistant registration systems:

First, all frequencies generally contribute equally, therefore, narrowly-banded noise is more easily handled in frequency domain.

25 Second, image data obtained under different illumination usually show slow-varying difference. Illumination non-uniformity usually appears as low-frequency variation across the image.

Also, carrier drifting in frequency domain, i.e.
30 phase tilt in spatial domain is low frequency.

stage tiling, slow change in stage height, and process variation are mostly low frequency noise. A/C noise is generally low frequency. Out-of-focus dusts are also at the lower side in the frequency domain. Back-reflection
5 noise is mostly relatively low frequency.

Random noise is typically at relatively high frequency. Both low frequency noise and high frequency noise are harmful to any mutual similarity measure and coherent peak convergence.

10 High frequency contents are independent of contrast reversal. A frequency-based technique is relatively scene independent and multi-sensor capable since it is insensitive to changes in spectral energy. Only frequency phase information is used for correlation, which is
15 equivalent to whitening of each image and whitening is invariant to linear changes in brightness and makes correlation measure independent.

Cross correlation is optimal if there is white noise. Therefore, a generalized weighting function can be
20 introduced into phase difference before taking the inverse Fourier transform. The weighting function can be chosen based on the type of noise immunity desired. So there are a family of correlation techniques, including phase correlation and conventional cross correlation.

25 For these reasons, the feature space can use prominent edges, contours of intrinsic structures, salient features, etc. Edges characterize object boundaries and are therefore useful for image matching and registration. The following are several candidate
30 filters to extract these features.

A Butterworth low pass filter is used to construct the BPF as follows:

$$weight = \frac{1}{1 + \left(\frac{r}{cutoff_2}\right)^{2 \cdot order}} - \frac{1}{1 + \left(\frac{r}{cutoff_1}\right)^{2 \cdot order}},$$

5

where *order* is the Butterworth order; *r* is the distance to DC; *cutoff₁* and *cutoff₂* are the cutoff frequencies at low end and high end respectively; and *weight* is the filter coefficient for the point.

10

The BPF can be used to choose any narrow band of frequency.

Edge Enhancement Filters in the Spatial Domain

15

Edge enhancement filters are used to capture information in edges, contours, and salient features. Edge points can be thought of as pixel locations of abrupt gray-level change. For a continuous image $f(x,y)$, its derivative assumes a local maximum in the direction of the edge. Therefore, one edge detection technique is to measure the gradient of f along r in a direction θ . The maximum value of $\partial f / \partial r$ is obtained when $(\partial / \partial \theta)(\partial f / \partial \theta) = 0$. This gives:

25

$$\left(\frac{\partial f}{\partial r}\right)_{\max} = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}, \text{ and}$$

$$\theta_{\text{edge}} = \arctan\left(\frac{\partial f}{\partial y} / \frac{\partial f}{\partial x}\right)$$

They can be re-written in digital form,

$$g(x,y) = g_x^2(x,y) + g_y^2(x,y) \quad , \text{ and}$$

5

$$\theta(x,y) = \arctan\left(\frac{g_y(x,y)}{g_x(x,y)}\right)$$

where $g_x(x,y)$ and $g_y(x,y)$ are orthogonal gradients along X and Y directions, obtained by convolving the image with a
10 gradient operator. To save computations, the magnitude gradient is often used

$$g(x,y) = |g_x(x,y)| + |g_y(x,y)|$$

15 The following lists some of the common gradient operators.

$$\text{Gradient } H_x = \begin{bmatrix} -1 & 1 \end{bmatrix} \quad H_y = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$$\text{Roberts } H_x = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad H_y = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

20

$$\text{Sobel } H_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad H_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

The first order derivative operators work best when the gray-level transition is quite abrupt, like a step
25 function. As the transition region gets wider, it is more

advantageous to apply the second-order derivatives.
Besides, these operators require multiple filter passes,
one in each primary direction. This directional
dependence can be eliminated by using the second-order
5 derivative operators. In some embodiments, a direction-
independent Laplacian filter is preferred and defined as

$$\nabla^2 f = \frac{\partial^2 f}{\partial^2 x} + \frac{\partial^2 f}{\partial^2 y}$$

10 The typical filter H has the form

$$H = \begin{bmatrix} -1 & -1 & -1 \\ -1 & C & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

where C is a parameter that controls the contents. The
15 value C=8 creates an edge-only filter, and sharp edges in
the original appear as a pair of peaks in the filtered
image. Values of C greater than 8 combine the edges with
the image itself in different proportions, and thereby
create an edge enhancement image.

20 In some instances, to increase correlation peak
height, it is also desirable to thicken the edges.
However, this process also broadens correlation peak and
hence reduces registration accuracy. It is maybe useful
for low-resolution match in a multi-resolution scheme.

25 In general, the purposes of edge enhancement filter
in spatial domain are: (1) to control information
contents to enter the registration flow; (2) to transform
the feature space; (3) to capture edge information of
salient features; (4) to sharpen correlation peak of
30 signal; (5) to solve the intensity reversal problem; and

(6) to have broader boundaries than edge detection or first derivative.

Thresholding in the Spatial Domain

5 The edge enhanced image still typically contains noise. However, the noise appears much weaker in the edge strength than intrinsic structures, and therefore, the edge-enhanced features can further be thresholded to remove points with small edge strength. In some
10 embodiments thresholding the filtered image can eliminate most of the A/C noise, D/C noise, and random noise.

 Threshold may be selected automatically by computing the standard deviation, σ , of the filtered image and using it to determine where the noise can be optimally
15 removed and there is still sufficient signal left for correlation. The threshold is defined as

$$threshold = numSigma \cdot \sigma$$

20 where numSigma is a parameter that controls the information contents entering the registration system. This parameter is preferably set up empirically.

 After thresholding, the points below the threshold are preferably disabled by zeroing them out, while the
25 rest of the points with strong edge strength are able to pass the filter and enter the following correlation operation. Notably, the idea of edge enhancement to boost the robustness and reliability of area-based registration is from the feature-based techniques. However, unlike the
30 feature-based techniques, the image is not thresholded to

a binary image. The filtered image is still gray-scale data by keeping the edge strength values of these strong edge points. The advantage of doing this is that the edge strength values of different edge points carry the
5 locality information of edges. The different locality information will vote differently in the correlation process. Therefore, this technique preserves the registration accuracy.

10 **Confidence of Image Matching**

This discussion is about correlation surface and the coherent peaks on the surface. As used in this discussion, features are the features of dominance, i.e. the major features in the scene. There are two types of
15 peaks on a correlation surface: coherent peaks and incoherent peaks. All peaks corresponding to features are coherent; all other peaks are incoherent, i.e. corresponding to noise.

Some examples of coherent peaks are as follows:

- 20 • Periodic signals with periods T_x and T_y in X and Y produce multiple periodic coherent peaks with the same periods. These peaks have approximately equal strengths, with the highest most likely at the center and peaks with fading strengths towards the
25 edge.
- Any locally repetitive signals also produce multiple coherent peaks. The highest coherent peak is most likely at the point of registration and all other secondary peaks are corresponding to local feature
30 repetitiveness.

- In many cases, correlation surface exhibits the behavior of a Sinc function typically seen as the response characteristics due to finite size of discrete Fourier transform in a system with limited bandwidth. The main lobe has the highest peak where the algorithm should converge at, but there are also multiple secondary lobes with peaks.

Incoherent peaks occur when noise exists. Random noise power is distributed randomly in some coherent peaks. Both A/C and D/C noises will bias, distort, and diverge the coherent peaks. Noise will also peel, fork, blur the coherent peaks.

The amplitude of the coherent peak is a direct measure of the congruence between the two images. More precisely, the power in the coherent peak corresponds to the percentage of dominant features in overlapping areas, while the power in incoherent peaks correspond to the percentage of noise and non-overlapping areas.

Therefore, the following two metrics are developed and used together to evaluate the quality of an image matching: First, the height of the first coherent peak. Second, the difference in strength, i.e. correlation coefficient, between the first coherent peak and the second peak, either coherent or incoherent.

An additional advantage using these metrics is that they are computed based on the correlation surface that is already available in real-time while computing alignment differences. The efficiency and real-time speed are critical in most image matching application where a

real-time confidence feedback signal is key to a successful automated target search systems such as wafer rotational alignment where an automated multi-FOV search is required.

5

Search Space and Subpixel Modeling

The task of the search strategy is often trivial in this implementation of registration since the whole correlation surface is already available, after inverse
10 Fourier transform, for searching. The point of registration is the maximum peak of magnitude correlation surface. One scan for the peak across the entire search space is typically sufficient. This is the integer registration detected.

15 To find out the subpixel offsets, a subpixel modeling is done as follows. A 2D parabola surface can be defined as

$$Z = ax^2 + by^2 + cxy + dx + ey + f.$$

20

This second-order polynomial is fit to 3x3-point correlation surface around the integer peak at (0,0).

$$\begin{bmatrix} Z_1 \\ Z_2 \\ \dots \\ Z_9 \end{bmatrix} = \begin{bmatrix} x_1^2 & y_1^2 & x_1 y_1 & x_1 & y_1 & 1 \\ x_2^2 & y_2^2 & x_2 y_2 & x_2 & y_2 & 1 \\ \dots & \dots & \dots & \dots & \dots & 1 \\ x_9^2 & y_9^2 & x_9 y_9 & x_9 & y_9 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \end{bmatrix}$$

25

where (x,y)'s are the coordinates of these 9 points,
which can be simplified to [-1, 0, 1] for both x and y. A
least-squared solution to the equation above based on
matrix pseudo-inverse operation gives an estimate for the
5 coefficients:

$$\begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \end{bmatrix} = \begin{bmatrix} x_1^2 & y_1^2 & x_1 y_1 & x_1 & y_1 & 1 \\ x_2^2 & y_2^2 & x_2 y_2 & x_2 & y_2 & 1 \\ \dots & \dots & \dots & \dots & \dots & 1 \\ x_9^2 & y_9^2 & x_9 y_9 & x_9 & y_9 & 1 \end{bmatrix}^{-1} \begin{bmatrix} Z_1 \\ Z_2 \\ \dots \\ Z_9 \end{bmatrix}$$

The subpixel locations of registration within this
10 3x3 block are found at the peak locations of the
parabola, which are determined by taking the partial
derivatives of the parabola equation with respect to x
and y and set them to zeros

$$\begin{aligned} \frac{\partial Z}{\partial x} &= 2ax + cy + d = 0 \\ \frac{\partial Z}{\partial y} &= 2by + cx + e = 0 \end{aligned}$$

15

which gives

$$x = \frac{2db - ce}{c^2 - 4ab}$$

$$y = \frac{2ae - dc}{c^2 - 4ab}$$

The coordinates of the integer peak and subpixel offsets are used to determine final registration offsets of the whole image.

FIGURE 1 shows an implementation of an intensity
5 based registration method. The method begins with
providing test intensity image 10 (which may also be
referred to as a first image) and reference intensity
image 12. Both images are separately edge enhanced 14
and 16 and then noise is removed from the edge enhanced
10 images using thresholding operations 18 and 20. The
images are then transformed 22 and 24 using a Fourier
transform.

The two transformed images are then used to in
coherence function computation 26 and an inverse Fourier
15 transform is applied thereto 28. Next a magnitude
operation is performed within a selected search range 30.
A confidence computation is then performed 32 and the
match of the images may then be either accepted or
rejected 34 based on the confidence value derived
20 therefrom. If the confidence value is within an
acceptable range, the registration process proceeds to
integer translation and subpixel modeling 36 and the
match of the images is accepted 38. If the confidence
value is not within an acceptable range, a new search is
25 initiated 40.

FIGURE 2 shows an implementation of a magnitude
based registration method. The method begins with
providing test hologram 50 and reference hologram 52.
Both holograms are separately transformed using a Fourier
30 transform 54 and 56 and a sideband extraction is applied

to each image 58 and 60. Next both image are separately
filtered with a bandpass filter 62 and 64. The resulting
images are then separately transformed using an inverse
Fourier transform 66 and 68 and a magnitude operation is
5 performed on each resulting image 70 and 72. The results
are then thresholded 74 and 76 before being transformed
using an Fourier transform operation 78 and 80.

The two transformed images are then used in
coherence function computation 82 and an inverse Fourier
10 transform is applied thereto 84. Next a magnitude
operation is performed within a selected search range 86.
A confidence computation is then performed 88 and the
match of the images may then be either accepted or
rejected 90 based on the confidence value derived
15 therefrom. If the confidence value is within an
acceptable range, the registration process proceeds to
integer translation and subpixel modeling 92 and the
match of the images is accepted 94. If the confidence
value is not within an acceptable range, a new search is
20 initiated 96.

FIGURE 3 shows an implementation of a phase image
based registration method. The method begins with
providing test hologram 100 and reference hologram 102.
Both holograms are separately transformed using a Fourier
25 transform 104 and 106 and a sideband extraction is
applied to each image 108 and 110. Next, both image are
separately filtered with a lowpass filter 112 and 114.
The resulting images are then separately transformed
using an inverse Fourier transform 116 and 118 and a
30 phase operation is performed on each resulting image 120

and 122. A phase-aware enhancement is then performed on the resulting images 124 and 126. The results are then thresholded 128 and 130 before being transformed using an Fourier transform operation 132 and 134.

5 The two transformed images are then used in coherence function computation 136 and an inverse Fourier transform is applied thereto 138. Next a magnitude operation is performed within a selected search range 140. A confidence computation is then performed 142 and
10 the match of the images may then be either accepted or rejected 144 based on the confidence value derived therefrom. If the confidence value is within an acceptable range, the registration process proceeds to integer translation and subpixel modeling 146 and the
15 match of the images is accepted 148. If the confidence value is not within an acceptable range, a new search is initiated 150.

FIGURE 4 shows an implementation of a complex based registration method. The method begins with providing
20 test hologram 152 and reference hologram 154. Both holograms are separately transformed using a Fourier transform 156 and 158 and a sideband extraction is applied to each image 160 and 162. The resulting images are then filtered using a bandpass filter 164 and 166.

25 The two filtered images are then used in coherence function computation 168 and an inverse Fourier transform is applied thereto 170. Next a magnitude operation is performed within a selected search range 172. A confidence computation is then performed 174 and the
30 match of the images may then be either accepted or

ATTORNEY DOCKET
068062.0166

25

rejected 176 based on the confidence value derived therefrom. If the confidence value is within an acceptable range, the registration process proceeds to integer translation and subpixel modeling 178 and the match of the images is accepted 180. If the confidence value is not within an acceptable range, a new search is initiated 182.

In some embodiments, simplification may be brought by eliminating confidence evaluation. The generally includes: (1) replacing coherence function computation with image conjugate product, i.e. without normalizing the cross power spectral density by maximum possible power of the two images, and (2) eliminating confidence computation and acceptance/rejection testing. The rest of the methods are essentially the same as in their original versions. As an example, the simplified version of a complex-based registration system is shown in FIGURE 5.

FIGURE 5 shows an simplified implementation of a complex based registration method. The method begins with providing test hologram 200 and reference hologram 202. Both holograms are separately transformed using a Fourier transform 204 and 206 and a sideband extraction is applied to each image 208 and 210. The resulting images are then filtered using a bandpass filter 212 and 214.

The two filtered images are then used to determine the image conjugate product 216 and an inverse Fourier transform is applied thereto 218. Next a magnitude operation is performed within a selected search range 220. The registration process proceeds to integer

ATTORNEY DOCKET
068062.0166

26

translation and subpixel modeling 222 and the match of the images is accepted and reported 224

Selection of a technique or a combination of multiple techniques for a specific application is a system engineering choice and depends on many factors. Among the important factors are basic functionality required, system optimization as a whole, data stream available, convenience and feasibility of filtering implementation, result of noise filtering and robustness, overall system speed and cost, system reliability.

The following examples are given to illustrate these principles.

Runtime Defect Detection

In the application of runtime wafer inspection, system speed and accuracy are essential. For this reason, the already available complex frequency data streams can be used to the advantage. Therefore, the registration may be simplified as shown in FIGURE 6.

FIGURE 6 shows a simplified implementation of a method for registering holographic complex images when sidebands are available in the datastream. The method begins with providing a test sideband 250 and reference sideband 252. Both sidebands are separately using a bandpass filter 254 and 256.

The two filtered images are then used to determine the image conjugate product 258 and an inverse Fourier transform is applied thereto 260. Next a magnitude operation is performed within a selected search range 262. The registration process proceeds to integer

translation and subpixel modeling 264 and the match of
the images is accepted and reported 266.

5 **Wafer Center Detection** (or die zero or other point
positional refinement.)

FIGURE 7 shows how the registration process is
applied to the aligning a wafer coordinate system to the
stage coordinate system. Wafer 300 is placed on a chuck
and images are acquired at candidate locations that
10 potentially match a stored reference pattern. The
procedure provided below is performed on the images to
determine the offset (Δx 302, Δy 304) between the actual
location of the reference pattern and the assumed
location of the pattern. The second step is to repeat the
15 registration procedure to determine and correct the
rotational angle, θ 306, between the die grid axis and
the stage axis.

In a particular embodiment of this application, we
have to use the full version of the algorithm:
20

Registration(translations, confidence, image1,
image2, ...)

which registers two images (of complex frequency, complex
25 spatial, magnitude, phase, or intensity) by computing
their translational differences and returns a realtime
confidence measure telling if it is a successful match,
the following procedures are developed for Wafer Center
Detection and Rotational Angle Detection.

Given an image chip as a template (e.g. 256x256),
the following steps are performed:

- Step 1. take an FOV 308, imagel, at the current
position where the template is taken (assuming it is an
5 image segment with features close to the real wafer
center).
- Step 2. zero-pad the template to the size of imagel.
- Step 3. call Registration(translations, confidence,
imagel, padded template, ...).
- 10 Step 4. if (confidence.maxCorrlist >= T1 and
confidence.measure >= T2)
Step. Output translations and compute wafer
center.
- Step 5. extract an image chip of 256x256 from imagel
15 at the location based on translation detected in Step 4.
- Step 6. Repeat Step 3 using template and the image
chip extracted (perform 256x256 registration).
- Step 7. Repeat Step 4.
- Step 8. Perform a circular search 311 by taking an
20 FOV from its neighbors with P% overlap, go to Step 3.
- Step 9. Repeat Step 4, Step 5, and Step 6 until the
condition in Step 4 is satisfied or signal it is out of
the search range predefined.
- Step 10. If no match is found within the search
25 range, output a failure signal and handle the case.

The steps above utilize the four parameters: T1, T2,
numSigma, and P%. T1 is the minimum Coors correlation
coefficient; T2 is the minimum confidence value; numSigma
30 is a noise threshold which controls the information

contents entering the registration system after edge enhancement; and P% is the overlap when taking an adjacent FOV. In one embodiment, in the case of zero-padding the template, the overlapped should be $\geq 50\% \times 256$ pixels since it only needs to cover a portion of the original template. Based on experiments, the following settings are typical for a successful search:

T1 = 0.4, T2 = 0.1, numSigma = 3.5.

10

Other parameters are similar to those in realtime registration.

In some embodiments the padding scheme can also be replaced with a tiling scheme.

15

ROTATIONAL ANGLE DETECTION

To identify rotational angle detection, given the wafer center, the following steps are performed:

Step 1. take an FOV 310, image1, along the wafer's center line on the left (this could also be the edge die for one-step alignment).

Step 2. take another FOV 312, image2, along the wafer's center line on the right, symmetric to the left FOV with respect to the wafer center.

Step 3. call Registration(translations, confidence, image1, image2, ...).

Step 4. if (confidence.maxCorr1st \geq T1 and confidence.measure \geq T2)

Stop. Output translations and compute rotational angle.

Step 5. Perform a spiral search by taking another FOV above or below with P% overlap, go to Step 3.

Step 6. Repeat Step 4 and Step 5 until the condition in Step 4 is satisfied or signal it is out of the search
5 range predefined.

Step 7. If no match is found within the search range, output a failure signal and handle the case.

The data should be taken along the wafer centerline detected above, or along a parallel line (where features
10 are guaranteed to present such as where template image is taken) close to the center to assure rotational accuracy.

The parameters are the same as in Wafer Center Detection. Note P% overlap in one direction (in case of spiral search, Y) will guarantee a $(50\% + P\%/2)$ overlap
15 area between a pair of FOVs in the worst case of gridding (gridding is where the data are actually taken with respect to the real location corresponding to its matching FOV).

The techniques described above provide a number
20 advantageous characteristics. Noise, including fixed-pattern (D/C noise), time-varying pattern (A/C noise), and random noise, may be removed up to 100% by a novel filter implemented in the spatial domain. This filter takes a different form for different data used.
25 Generally, it first enhances edges of high-frequency spatial features. Only strong features can pass the filter and noise is left out of the process. The gray-scale edge strength data, instead of raw intensity/phase, is then used in the following correlation process.

The correlation process is implemented in Fourier domain for speed and efficiency. In most embodiments a Fast Fourier Transform (FFT) is used to implement the Fourier transform operations.

5 The use of a confidence value for each match is advantageous. This confidence value is defined using the peak pattern of 2-D correlation surface. Together with correlation coefficient, this confidence value provides a reliable measure of the quality of image matching.

10 Providing a mechanism for a fully automated searching (in combination with a mechanical translation of the target object) from as many fields of view (FOVs) as required until the right target is matched is also advantageous. The quality of each move is gauged by a
15 confidence defined during registration computation process, and the confidence value can further be used to accept a match or reject it and initiate a new search.

Automated wafer rotational alignment fully automates the correction of any wafer rotational errors. This is
20 important for initial wafer setup in a wafer inspection system. It reduces setup time of operators and achieves the required accuracy for wafer navigation. The registration system provides the inspection system a robust, reliable, and efficient sub-system for wafer
25 alignment.

The methods described promote flexibility in accepting of a variety of input data. In case of DDH wafer rotational alignment, this method may accept five major data formats and compute registration parameters
30 directly based on these data: a. complex frequency data;

b. complex spatial data; c. amplitude data extracted from
a hologram; d. phase data extracted from a hologram; and
e. intensity-only data. This flexibility provides
opportunities to develop more reliable and efficient
5 system as a whole.

Comparing Holographic Images

The present invention also includes systems and
methods for comparing holographic images for the purpose
10 of identifying changes in or differences between objects.
As shown in FIGURE 8, the imaging system, depicted
generally at 340 includes the primary components: 1)
mechanical positioning system 380 with computer control
linked to a system control computer 350; 2) optical
15 system 370 for creating a hologram including an
illumination source; 3) data acquisition and processing
computer system 360; 4) processing algorithms operable to
execute on processing system 360 and may also include 5)
a system for supervisory control of the subsystems (not
20 expressly shown).

Imaging system 340 operates by positioning, in up to
six degrees of freedom (x, y, theta, z, tip, tilt) one
instance of an object in the field of view (FOV) of the
optical system and acquiring a digital hologram using
25 acquisition system 360 and performing the first stage of
hologram processing. The resulting intermediate
representation of the image wave may be stored in a
temporary buffer.

Positioning system 380 is then instructed to move to
30 a new location with a new object in the FOV and the

initial acquisition sequence is repeated. The coordinates that the positioning system uses for the new location is derived from a virtual map and inspection plan. This step and acquire sequence is repeated until a second instance
5 of the first object is reached.

A distance-measuring device is preferably used in combination with positioning system 380 to generate a set of discrete samples representative of the distance between the object and the measuring device. A
10 mathematical algorithm is then used to generate a map with a look-up capability for determining the target values for up to three degrees of freedom (z, tip, tilt) given as input up to three input coordinates (x, y, theta).

At this point optics system 370 acquires the hologram of the second instance of the object and it is processed to generate an intermediate representation of the image wave. The corresponding representation of the first instance is retrieved from the temporary buffer and
20 the two representations are aligned and filtered. Many benefits can be realized at this point by performing unique processing on the representation of the object in the frequency domain. A comparison (reference difference image description) between these two instances may be
25 made and the result stored in a temporary buffer. This process may be repeated for additional FOVs containing second instances of the objects.

Positioning system 380 reaches a third instance of the object and the two previous steps (intermediate
30 representation and comparison to second instance) are

completed. The results of the comparison between the first and second instance is retrieved from the temporary buffer and a noise suppression and source logic algorithm may preferably be applied to the retrieved and current
5 comparisons.

The results may then be analyzed and summary statistics generated. These results are conveyed to the supervisory controller. This cycle is repeated as new instances of the objects are acquired.

10

Generating the Difference between Complex Images

The present invention contemplates variations for generating the difference between two complex images.

An amplitude difference may be utilized. First,
15 both complex images are preferably converted to an amplitude representation, and the magnitude of the difference between the resulting amplitudes (pixelwise) is computed. In one embodiment, this represents the difference in reflectance between the two surfaces being
20 imaged.

A phase difference may be utilized. First both complex images are preferably converted to a phase representation and the effective phase difference between the resulting phase values (pixelwise) is computed. This
25 may be performed directly as described, or by computing the phase of the pixelwise ratio of the two images after they have each been amplitude normalized. In one embodiment this represents a height difference between the two surfaces being imaged.

Also, a vector difference may be utilized. First the two complex images are subtracted directly in the complex domain, then the amplitude of the resulting complex difference is computed. This difference combines
5 aspects of the amplitude difference and phase difference in an advantageous way. For example, in situations where the phase difference is likely to be noisy, the amplitude is likely to be small, thus mitigating the effects of the phase noise on the resulting vector difference.

10

Aligning and Comparing Two Consecutive Difference Images

The present invention further contemplates the alignment and comparison of two consecutive difference images in order to determine which differences are common
15 to both. The amount to shift one difference image to match the other is typically known from earlier steps performed to compute the difference images originally; namely, image *A* is shifted by an amount *a* to match image *B* and generate difference image *AB*, while image *B* is
20 shifted by an amount *b* to match image *C* and generate difference image *BC*. The appropriate amount to shift image *BC* to match image *AB* is therefore $-b$. Three alternate approaches to determining which differences the two difference images have in common are described below.

25 In one embodiment, the difference images are thresholded, then one of the two thresholded images is shifted by the appropriate amount, rounded to the nearest whole pixel. The common differences are then represented by the logical-AND (or multiplication) of the shifted and
30 unshifted thresholded difference images.

In another embodiment the difference images are first shifted by the appropriate (subpixel) amount before thresholding and then the image is thresholded. The common differences are then computed by a logical-AND (or
5 multiplication) as above.

In another embodiment, one of the difference images is shifted by the appropriate (subpixel) amount and combined with the second image before thresholding. The combination of the two images can be any one of several
10 mathematical functions, including the pixelwise arithmetic mean and pixelwise geometric mean. After combining the two difference images, the result is then thresholded.

15 **Example Operations**

The discussion below provides a description of example operations of the present invention. First, a hologram is acquired with a CCD camera (as shown in FIGURES 9 and 10) and stored in memory. The object wave
20 is defined as

$$A(x, y)e^{i(\vec{k}_A \cdot \vec{r} + \phi_A(x, y))}$$

and the reference wave as

$$B(x, y)e^{i(\vec{k}_B \cdot \vec{r} + \phi_B(x, y))}.$$

25 The intensity of the recorded hologram, ignoring camera nonlinearities and noise, is:

$$I_{hol} = \left| A(x, y)e^{i(\vec{k}_A \cdot \vec{r} + \phi_A(x, y))} + B(x, y)e^{i(\vec{k}_B \cdot \vec{r} + \phi_B(x, y))} \right|^2 \quad (1)$$

The phase difference $\Delta\phi(\vec{r})$ between the two waves is defined as $\Delta\phi(\vec{r}) = (\phi_A - \phi_B)$ and the vector difference $\Delta\vec{k}$, which represents the angle between the two arms, as $\Delta\vec{k} = (\vec{k}_A - \vec{k}_B)$. Equation (1) simplifies to:

$$5 \quad I_{hol} = A^2(\vec{r}) + B^2(\vec{r}) + 2\mu_0 A(\vec{r})B(\vec{r})\cos(\Delta\vec{k}\vec{r} + \Delta\phi(\vec{r})) \quad (2)$$

where μ_0 represents the coherence factor. Edgar has documented further details along these lines.

In preferred embodiments this step may be implemented either as a direct image capture and transfer
10 to memory by a digital holographic imaging system itself, or simulated in an off-line program by reading the captured image from disk. In this particular preferred embodiment, the image is stored as 16-bit grayscale, but with 12 bits of actual range (0-4095) because that is the
15 full range of the camera.

Next, the holographic image is preferably processed to extract the complex wavefront returned from the object as shown in FIGURE 11. In one preferred embodiment, a Fast Fourier Transform (FFT) is performed on the captured
20 (and optionally enhanced) hologram. The FFT of the hologram intensity is expressed as:

$$\begin{aligned} FFT\{I_{hol}\} = & \delta_{(\vec{q}=0)} * FFT\{A^2(\vec{r}) + B^2(\vec{r})\} + \\ & \mu_0 \cdot \delta_{(\vec{q}=-\Delta\vec{k})} * FFT\{A(\vec{r})B(\vec{r})e^{i(\Delta\vec{k}\vec{r} + \Delta\phi)}\} + \\ & \mu_0 \cdot \delta_{(\vec{q}=\Delta\vec{k})} * FFT\{A(\vec{r})B(\vec{r})e^{i(\Delta\vec{k}\vec{r} + \Delta\phi)}\} \end{aligned} \quad (3)$$

Next, a carrier frequency of a holographic image is
25 found. In one embodiment, this first requires that the frequency where the sideband is centered, as shown in FIGURE 12, must be located in order to isolate the

sideband properly. This may either be done on the first
hologram processed and the same location used for all
subsequent images, or the carrier frequency can be
relocated for every single hologram. First the location
5 $\bar{q} = \Delta \bar{k}$ (or $\bar{q} = -\Delta \bar{k}$) from the hologram FFT in equation (2)
is sought. Because the modulus of the sidebands exhibits
peaks at these two locations, the desired location can be
found by searching the modulus of $FFT\{I_{hol}\}$ away from $\bar{q}=0$.

In some embodiments a search area for the sideband
10 is defined as a parameter. The modulus of the hologram
FFT is computed in the defined area, and the location of
the maximum point is chosen as the carrier frequency.
The search area may be specified as a region of interest
(maximum and minimum x and y values) in all
15 implementations.

In a particular embodiment, the carrier frequency is
computed to sub-pixel accuracy by interpolation of the
FFT modulus in the area of the found maximum. To correct
for the sub-pixel location of the carrier frequency, the
20 FFT is then modulated by a phase-only function after
isolating the sideband

The search area for the sideband, may be specified
either as a region of interest in the Fourier domain or
as the number of pixels away from the x and y axes not to
25 search in the Fourier domain. In some embodiments this
parameter may be selectively modified. Alternatively, a
user may optionally set the manual location of the
sideband, which sets the carrier frequency location to a
fixed value that is used for all images. (In the a

particular embodiment, the same effect can be achieved by setting the search area to be a single point.)

For an inspection series, the carrier frequency may be assumed to be stable and therefore need not be recomputed for each hologram. The carrier frequency can be found once and that frequency used for all subsequent holograms during the same inspection.

After the sideband is located, a quadrant of the hologram FFT centered at the carrier frequency is extracted as shown in FIGURE 13. This isolation of the sideband quadrant takes one of the sideband terms from equation (3) and modulates it to remove the dependence on $\Delta k \vec{r}$:

$$sideband \equiv \mu_0 \cdot FFT\{A(\vec{r})B(\vec{r})e^{i\Delta\phi}\} \quad (4)$$

Implementation of this step is straightforward. Note that in some embodiments, a quadrant is not extracted from the FFT, but rather the FFT is recentered at the carrier frequency and left at its original resolution.

The extracted sideband may then be filtered. In a particular embodiment, a Butterworth lowpass filter is applied to the extracted sideband to reduce the effect of any aliasing from the autocorrelation band and to reduce noise in the image.

The lowpass filter $H(\vec{q})$ is applied to the sideband as shown in FIGURE 14. The filtered sideband is the FFT of the complex image wave $\psi(\vec{r})$ that we wish to reconstruct:

$$FFT\{\psi(\vec{r})\} = sideband \cdot H(\vec{q}) \quad (5)$$

The Butterworth lowpass filter is defined by the equation:

$$H(\vec{q}) = \frac{1}{1 + \left(\frac{|\vec{q}|}{q_c} \right)^{2N}} \quad (6)$$

where q_c is the cutoff frequency of the filter (that is, the distance from the filter center where the gain of the filter is down to half its value at $|\vec{q}|=0$) and N is the order of the filter (that is, how quickly the filter cuts off).

In embodiments where off-axis illumination is used, the lowpass filter may need to be moved off-center to capture the sideband information more accurately. Letting \vec{q}_{off} represent the location where we wish to place the center of the filter (the offset vector), the equation for the Butterworth filter is:

$$H(\vec{q}) = \frac{1}{1 + \left(\frac{|\vec{q} - \vec{q}_{off}|}{q_c} \right)^{2N}} \quad (7)$$

In preferred embodiments the Butterworth filter should be computed only once for the given parameters and image size and stored for use with each image.

In preferred embodiments the cutoff frequency, also called the filter "size" or "radius", and order of the filter must be specified.

If an off-axis filter is desired, the offset vector for the center of the filter should also be specified; this parameter should also be selectively adjustable. In a preferred embodiment a flag indicating whether to use a lowpass filter or bandpass filter may allow a use to

select the type of filter employed in the processing software.

In some embodiments, processing software programs have the ability to substitute a bandpass filter for the
5 lowpass filter. Using the bandpass filter has been shown to improve defect detection performance on particular defect wafers. The bandpass filter is implemented as a series multiplication of Butterworth lowpass and highpass filters; the highpass filter may be defined as "one minus
10 a lowpass filter" and has the same type of parameters to specify as the lowpass filter.

Next the inverse Fast Fourier Transform (IFFT) is performed on the filtered sideband to derive the complex image wave to produce a magnitude image and phase image
15 as shown FIGURES 15 and 16, respectively. The IFFT of the filtered sideband yields:

$$\psi(\vec{r}) = \mu_0 A(\vec{r}) B(\vec{r}) e^{i\psi(\vec{r})} \quad (8)$$

where it has been assumed that the aperture of the lowpass filter perfectly isolates the sideband. In
20 practice, this is not possible, but the assumption is necessary to achieve a tractable expression, and equation (7) does represent the results reasonably well.

If the phase of the resulting complex image is not flat enough (i.e., there are several phase wraps across
25 the image), flat field correction may be applied to improve the results. This consists of dividing the complex image by the complex image of a reference flat (mirror) to correct for variations in illumination intensity and (especially) background phase.

First, $\phi(\vec{r})$ represents the complex image of a reference flat hologram (processed as described above). The flat field corrected hologram $\psi'(\vec{r})$ is:

$$\psi'(\vec{r}) = \frac{\psi(\vec{r})}{\phi(\vec{r})} \quad (9)$$

5 To implement this step, during a prior inspection run, a flat field hologram is processed to a complex image. That image is stored and divided pixelwise into each complex image from the run. Typically, the parameters used to generate complex images (sideband
10 search area and filter parameters) are the same for the flat field hologram as for the inspection holograms.

The reference flat corrects for intensity as well as phase, and as a result modulus images $|\psi'(\vec{r})|$ resulting from equation (8) may not be very useful for viewing or
15 magnitude only processing algorithms. This problem can be alleviated by modifying the reference flat image $\phi(\vec{r})$ to have unit modulus at every pixel. The flat field correction then only corrects for non-flat phase in the inspection images.

20

Differencing Operations

Differencing operations are necessary to identify difference between two corresponding complex images. One preferred method of performing differencing operation is
25 outlined below.

After obtaining two complex images, the two images are aligned so that a direct subtraction of the two images will reveal any differences between the two. In this embodiment the registration algorithm is based on

the cross-correlation of the two images. Because the registration algorithm is based on the cross-correlation of the two images, performance may be improved by removing the DC level and low-frequency variation from the images. This allows the high-frequency content of sharp edges and features to be more prominent than any alignment of low-frequency variations.

A Butterworth highpass filter $H_{HP}(\vec{q})$ may be applied (in the frequency domain) to each of the complex images ψ_1 and ψ_2 to be registered:

$$\Psi_n(\vec{q}) = FFT\{\psi_n\} = sideband \cdot H(\vec{q}) \cdot H_{HP}(\vec{q}) \quad (10)$$

This effectively bandpass filters the images. The highpass filter H_{HP} is defined as:

$$H_{HP}(\vec{q}) = 1 - \frac{1}{1 + (|\vec{q}|/q_c)^{2N}} \quad (11)$$

15

Implementation of the highpass filtering step is straightforward. The size of the highpass filter used can be user-defined or determined as a fixed percentage of the size of the lowpass filter applied in above. The highpass filter is preferably computed once and stored for application to every image.

The cutoff frequency and order of the highpass filter H_{HP} may specified by the user or fixed to a pre-defined relationship with the lowpass filter parameters. In some embodiments it may be desirable to limit the parameters of this step to a fixed relationship with the lowpass filter parameters in order to reduce the number of user variables.

After filtering, the cross-correlation of the two images is computed. The peak of the cross-correlation surface preferably occurs at the location of the correct registration offset between the images.

5 The cross-correlation $\gamma_{n,n+1}(\bar{r})$ between the two bandpass filtered images is computed by taking the inverse Fourier transform of the product of the first image with the conjugate of the second image:

$$\gamma_{n,n+1}(\bar{r}) = IFFT\{\Psi_n(\bar{q}) \cdot \Psi_{n+1}^*(\bar{q})\} \quad (12)$$

10

The registration offset between the two images corresponds to the location where the cross-correlation surface achieves its maximum. The registration offset between the two images is the value of \bar{r} , denoted \bar{r}_{\max} ,
15 for which $|\gamma(\bar{r})|$ is a maximum. A region centered at the origin of the cross-correlation is searched for the maximum value. Once the location of the maximum is found, a quadratic surface is fit to the 3x3 neighborhood centered at that location, and the subpixel location of
20 the peak of the fit surface is used as the subpixel registration offset. The equation for the quadratic surface is:

$$ax^2 + bxy + cy^2 + dx + ey + f = 0 \quad (13)$$

The values of the coefficients a , b , c , d , e , and f
25 are calculated via a matrix solve routine. A 9x6 matrix (A) of values in the 3x3 neighborhood for the terms x^2 , xy , etc. is calculated, and the 6x1 vector of (unknown) coefficients $\bar{z} = [a \ b \ c \ d \ e \ f]^T$ is formed. The values of the cross-correlation corresponding to each location are put

into a 9x1 vector $\vec{h} = [y(\vec{r}_1) \ y(\vec{r}_2) \ \dots]^T$. The form of the matrix A is:

$$A = \begin{bmatrix} x_1^2 & x_1 y_1 & y_1^2 & x_1 & y_1 & 1 \\ x_2^2 & x_2 y_2 & y_2^2 & x_2 & y_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_9^2 & x_9 y_9 & y_9^2 & x_9 & y_9 & 1 \end{bmatrix} \quad (14)$$

- 5 The coefficients of the fitted surface are then found by solving equation (14) for \vec{z} :

$$A \cdot \vec{z} = \vec{h} \quad (15)$$

- 10 The location of the maximum of the quadratic surface (x_{\max}, y_{\max}) is then computed from the coefficients \vec{z} and used as the subpixel registration offset value.

$$x_{\max} = \frac{2c \cdot d - b \cdot e}{b^2 - 4a \cdot c} \quad (16)$$

$$y_{\max} = \frac{2a \cdot e - b \cdot d}{b^2 - 4a \cdot c} \quad (17)$$

- 15 The determination of the location where the cross-correlation surface is maximum can be achieved in several different ways. In one implementation, the interpolation may be performed by fitting a quadratic surface to the 3x3 neighborhood centered at the maximum, and finding the
20 location of the maximum of the fitted surface. In another implementation, there is an option to perform this interpolation using three points in each (x and y) direction separately.

- Typically the maximum registration offset must be
25 specified, usually as a maximum number of pixels in any

direction the images may be shifted relative to each other to achieve alignment.

The registration shift determination described essentially completes the registration process. Note
5 that this process generally corresponds with the more registration process described in greater detail above.

After determining the registration shift between the two images, the first image is shifted by that amount to align it to the second image. The image $\psi_n(\tilde{r})$ is shifted
10 by the registration amount \tilde{r}_{\max} :

$$\psi'_n(\tilde{r}) = \psi_n(\tilde{r} - \tilde{r}_{\max}) \quad (18)$$

Because the registration shift \tilde{r}_{\max} is typically a non-integer value, a method of interpolating the sampled
15 image must be chosen. The two preferred methods for interpolation are bilinear interpolation and frequency domain interpolation. Bilinear interpolation works in the spatial domain using the four nearest whole pixels to the desired subpixel location. Assume we wish to find
20 the interpolated value of ψ at the location $(x+\Delta x, y+\Delta y)$, where x and y are integers and $0 \leq x < 1$ and $0 \leq y < 1$. The bilinearly interpolated value is computed as:

$$\begin{aligned} \psi(x+\Delta x, y+\Delta y) = & (1-\Delta x) \cdot [(1-\Delta y) \cdot \psi(x, y) + \Delta y \cdot \psi(x, y+1)] \\ & + \Delta x \cdot [(1-\Delta y) \cdot \psi(x+1, y) + \Delta y \cdot \psi(x+1, y+1)] \end{aligned} \quad (19)$$

25 Frequency domain interpolation is performed using a basic shifting property of the Fourier transform:

$$\psi(x+\Delta x, y+\Delta y) = IFFT\{\Psi(u, v) \cdot e^{-j2\pi(\Delta x u + \Delta y v)}\} \quad (20)$$

For equation (19), the range of Δx and Δy is not limited.

The two images being compared must be normalized so that when subtracted their magnitude and phase will align and yield near-zero results except at defects. There are two major methods used to normalize the complex images. In the first and simplest method, termed "complex normalization," the first image of the pair is normalized to the second by multiplying it by the ratio of the complex means of the two images. The complex mean of an image is defined as:

$$\mu_{\psi} = \frac{1}{N^2} \left(\sum_{\vec{r}} \text{Re}\{\psi(\vec{r})\} + i \cdot \sum_{\vec{r}} \text{Im}\{\psi(\vec{r})\} \right) \quad (21)$$

where N^2 is the number of pixels in the image. The equation to normalize the image $\psi'_n(\vec{r})$ to $\psi_{n+1}(\vec{r})$ is:

$$\psi'_n(\vec{r}) = \frac{\mu_{\psi_{n+1}}}{\mu_{\psi_n}} \cdot \psi'_n(\vec{r})$$

In the second method, termed "magnitude-phase normalization," the magnitude and phase of the images is aligned directly, rather than the real and imaginary parts. First, the means of the image magnitudes are computed:

$$\mu_{|\psi|} = \frac{1}{N^2} \sum_{\vec{r}} |\psi(\vec{r})| \quad (22)$$

Second, the phase offset between the two images is computed. The phase difference between the two images is computed:

$$\angle \psi_{n+1} - \angle \psi'_n = \angle \left(\frac{\psi_{n+1}}{\psi'_n} \right) = \tan^{-1} \left(\frac{\psi_{n+1}}{\psi'_n} \right)$$

To find the phase offset, we need to compute the phase shift of this phase difference image that will yield the fewest phase jumps in the image. Because this image is expected to be somewhat uniform, it is more
5 reliable to find the phase offset that results in the greatest number of phase jumps, and then posit that correct phase offset is π radians offset from that. The result is a phase offset $\Delta\phi$ that will be used with the magnitude mean ratio to normalize the first image to the
10 second:

$$\psi'_{n+1}(\vec{r}) = \frac{\mu_{|\psi_{n+1}|}}{\mu_{|\psi'_n|}} \cdot \psi'_n(\vec{r}) \cdot e^{-i\Delta\phi} \quad (23)$$

Implementation of this step is fairly straightforward from the mathematical description. The
15 magnitude-phase normalization is often more computationally intense and may be unnecessary if the wavefront matching step is used. If wavefront matching is used, it is not necessary to perform the normalization step at all, because the wavefront matching is a form of
20 normalization.

Wavefront matching adjusts the phase of the second image by a filtered version of the phase ratio between the images, in order to remove low-frequency variation from the difference image cause by phase anomalies.
25 First, the phase difference between the images is found by dividing the two complex images:

$$\rho(\vec{r}) = \frac{\psi'_n(\vec{r})}{\psi_{n+1}(\vec{r})} \quad (24)$$

This ratio is then lowpass filtered in the frequency domain using a filter with a very low cutoff frequency:

$$\rho_{fil}(\vec{r}) = IFFT\{FFT\{\rho(\vec{r})\} \cdot L(\vec{r})\} \quad (25)$$

where $L(\vec{r})$ is a third-order Butterworth lowpass

5 filter with a cutoff frequency of six pixels. This filtered ratio is used to modify the second image so that low frequency variations in the phase difference are minimized:

$$\tilde{\psi}_{n+1}(\vec{r}) = \psi_{n+1}(\vec{r}) \cdot \rho_{fil}(\vec{r}) \quad (26)$$

10

Implementation of this step is straightforward using the above equations. The order and cutoff frequency of the lowpass filter used in this step are fixed. Also, note that in one preferred embodiment the second image is 15 the one modified by this algorithm, not the first. This is to minimize the number of pixels where the ratio $\rho(\vec{r})$ will be undefined because of zeroes in the denominator at border pixels.

In some instances the differences among 20 implementations for handling border pixels when shifting images may cause this step to propagate differences throughout the images. Until and unless the handling of border pixels during shifting is identical among the various implementations, the wavefront matching step will 25 result in differences throughout the images. Typically these differences are quite small. Also, wavefront matching can cause artifacts near the borders because of the periodicity assumption of the FFT. The effects of these artifacts can extend beyond the border region 30 excluded from defects.

The vector difference between the two registered, normalized, phase corrected images is then computed as shown in the first difference image shown in FIGURE 17 and the second difference image as shown in FIGURE 18.

5 The vector difference between the images is

$$|\Delta\psi_{n+1,n}(\vec{r})| = |\tilde{\psi}_{n+1}(\vec{r}) - \psi_n^*(\vec{r})| \quad (27)$$

The implementation of this step is straightforward. Note that in alternate embodiments phase differences and magnitude differences may also be used to detect defects.

10 Pixels near the edges of the difference image are set to zero to preclude defect detection in those areas, which are prone to artifacts. Each pixel in the vector difference image that is within a specified number of pixels of each edge of the image is set to zero. This requires that the number of pixels at each edge to zero out must be specified. In some embodiments, the number of pixels is taken to be equal to the maximum allowed registration shift, in pixels

20

Defect Detection

The vector difference image is thresholded to indicate the location of possible defects between each pair of images as shown in FIGURES 19 and 20. The standard deviation σ of the vector difference image $|\Delta\psi_{n+1,n}(\vec{r})|$ is computed. A threshold is set at a user-specified multiple of the standard deviation, $k\sigma$, and the difference image is thresholded at that value:

$$\delta_{n+1,n}(\vec{r}) = \begin{cases} 1, & |\Delta\psi_{n+1,n}(\vec{r})| > k\sigma \\ 0, & |\Delta\psi_{n+1,n}(\vec{r})| \leq k\sigma \end{cases} \quad (28)$$

The initial threshold value is computed based on the standard deviation of the entire difference image. In the one implementation, the threshold is iteratively
5 modified by recomputing the standard deviation excluding pixels above the threshold until there are no further changes. This effectively lowers the threshold for images that have many defects, sometimes quite substantially. In a preferred embodiment the multiple of
10 the standard deviation at which to threshold the image is specified by the user.

The two thresholded difference images used to ascertain which image a defect originates from are then aligned. Because the first image of any pair is aligned
15 to the second image of the pair, the two resulting difference images are in different frames of reference. In a sequence of three complex images that are compared to each other, ψ_1 , ψ_2 , and ψ_3 , the first thresholded difference $\delta_{2,1}$ is aligned with ψ_2 , and the second
20 difference $\delta_{3,2}$ is aligned with ψ_3 . Since these two thresholded difference images will yield the defects for the image ψ_2 , the image $\delta_{3,2}$ must be shifted so that it is aligned with ψ_2 . Because of the binary nature of the thresholded images, it is only necessary to align the
25 images to whole-pixel accuracy. The registration shift between the images ψ_2 and ψ_3 is already known from earlier computations to sub-pixel accuracy; this shift,

\bar{r}_{\max} , is rounded to the nearest whole pixel (denoted \bar{r}'_{\max}) and applied to $\delta_{3,2}$ in the direction opposite to its earlier application (which shifted image 2 into alignment with image 3):

5
$$\delta'_{3,2}(\bar{r}) = \delta_{3,2}(\bar{r} + \bar{r}'_{\max}) \quad (29)$$

The implementation of this step is straightforward.

Next, a logical AND operation is applied to the aligned thresholded difference images to eliminate any
10 detected defects that do not appear in both as shown in FIGURE 21. This reduces the number of false positive defects and assigns defects to the proper image in the sequence.

The defects in the image ψ_2 found when it is
15 compared to the corresponding images ψ_1 and ψ_3 , are given by:

$$d_1(\bar{r}) = \delta_{2,1}(\bar{r}) \cap \delta'_{3,2}(\bar{r}) \quad (30)$$

In one particular embodiment the logical AND is implemented as a multiplication of the two thresholded
20 images, since their values are limited to be either 0 or 1.

In an alternate embodiment the above steps may be reordered, so that the alignment and logical AND steps are performed before thresholding, subpixel alignment may
25 be used instead, and the logical AND step becomes a true multiplication.

In some embodiments, the resulting defect areas may be disregarded if they fall below a certain size threshold. Also, morphological operations on the defect

areas may be used to "clean up" their shapes. Shape modification may be implemented as a mathematical morphology operation, namely the morphological closing. This operator is described as follows.

- 5 Let K denote the structuring element (or kernel) for the morphological operator. Define the symmetric set $\tilde{K} = \{-\tilde{r} : \tilde{r} \in K\}$, which is a reflection of K about the origin. The translation of a set to a point \tilde{s} is denoted by a subscript; for example, the set K translated to the
10 point \tilde{s} is $K_{\tilde{s}}$. The set processing morphological erosion and dilation are defined by:

$$\text{Erosion:} \quad d \ominus \tilde{K} = \{\tilde{s} : K_{\tilde{s}} \subseteq d(\tilde{s})\} = \bigcap_{\tilde{r} \in K} d(-\tilde{r}) \quad (31)$$

$$\text{Dilation:} \quad d \oplus \tilde{K} = \{\tilde{s} : (K_{\tilde{s}} \cap d(\tilde{s})) \neq \emptyset\} = \bigcup_{\tilde{r} \in K} d(\tilde{r}) \quad (32)$$

- The symbols \ominus and \oplus denote Minkowski subtraction and
15 Minkowski addition, respectively. The erosion of a binary image d has true pixels where the structuring element K may be translated while remaining entirely within the original area of true pixels. The dilation of d is true where K may be translated and still intersect
20 the true points of d at one or more points.

The morphological opening and closing operations are sequential applications of the erosion and dilation, as follows:

$$\text{Opening:} \quad d \circ K(\tilde{r}) = [(d \ominus \tilde{K}) \oplus K](\tilde{r}) \quad (33)$$

$$25 \text{ Closing:} \quad d \bullet K(\tilde{r}) = [(d \oplus K) \ominus K](\tilde{r}) \quad (34)$$

Morphological closing with a square kernel (K) is the most likely operation for shape modification of the defect map d .

Size restriction may be implement by counting the
5 number of pixels in each connected component. This step will likely be combined with connected component analysis. In one embodiment, shape modification utilizes a mathematical morphology operation, particularly morphological closing with a 3×3 square kernel.

10 In a preferred embodiment the minimum defect size to accept must be specified for the size restriction operation. In some embodiments this parameter may be user-modified. For shape modification operations the size and shape of the kernel plus the type of
15 morphological operator must be specified by the user. Additionally the user may also specify whether to use shape modification at all.

Areas with non-zero pixels in the resulting defect image $d_i(\vec{r})$ are converted to a "connected component"
20 description. The connected component routine preferably looks for defect clusters that are continuous in the x direction. Once a linear string of defects is identified, it is merged with other blobs it may be in contact with in the y direction. Merge involves
25 redefining the smallest bounding rectangle that completely encloses the defect cluster. A limit such as for 50 defects may be imposed in the detection routine to improve efficiency. If at any point, the defect label exceeds the limit plus a margin, the analysis is aborted.

Once the entire image is scanned, the merge procedure is repeated continuously until the defects do not increase.

5 The connected components are then shown as a magnitude image as shown in FIGURE 22 or a phase image as shown in FIGURE 23. In one embodiment the connected components are mapped into a results file and basic statistics for the defects are computed. In one particular embodiment only the coordinates of the bounding rectangles of the defects are reported.

10 Although the disclosed embodiments have been described in detail, it should be understood that various changes, substitutions and alterations can be made to the embodiments without departing from their spirit and scope.